Time-Explicit Graphs: A Framework for Dynamic Network Analysis

[Extended Poster-Abstract]

René Pfitzner
ETH Zurich
Chair of Systems Design
Zurich, Switzerland
rpfitzner@ethz.ch

Ingo Scholtes
ETH Zurich
Chair of Systems Design
Zurich, Switzerland
ischoltes@ethz.ch

Claudio J. Tessone
ETH Zurich
Chair of Systems Design
Zurich, Switzerland
tessonec@ethz.ch

Antonios Garas
ETH Zurich
Chair of Systems Design
Zurich, Switzerland
agaras@ethz.ch

Frank Schweitzer
ETH Zurich
Chair of Systems Design
Zurich, Switzerland
fschweitzer@ethz.ch

ABSTRACT

When dealing with networks of several kind, network theory provides the theoretical framework and the practical tools to analyze them. Mostly analysis focuses on steady networks, i.e. networks with fixed topology. However, real-world scenarios often deal with (highly) dynamic topologies. As of today, the question of how to investigate such scenarios, e.g. defining a meaningful notion of "temporal centrality", has generally not been answered. In this poster we present research pointing out the obstacles arising in this endeavor as well as a possible framework to address those issues, which we call *time-explicit graphs*.

General Terms

Dynamic Networks, Network Analysis, time-explicit graphs

1. THE NEED FOR DYNAMIC NETWORK ANALYSIS

The rise of network analysis over the last two decades impressively demonstrates the influence new tools and technologies can have on the way research is done. Often it is the unifying tool that leads to interdisciplinary exchange: from physics to sociology, from mathematics to biology. The usefulness and success of network analysis basically arises because of its rather intuitive and widely applicable framework. However, as of today network analysis deals (almost exclusively) with static topologies. In real-world examples, however, dynamic topologies are observed: routing networks, networks of social relationships or coauthorship networks are just a few examples. Very often the time dimension, and all information about the system encoded in this, is neglected and only the aggregated network is studied. Thus, questions like "Who is the most influential node?" are reduced to an aggregated picture, which might lead the researcher to false implications. This is most intuitively seen in aggregated graphs of social networks: in the aggregated graph someone who initiated a high number of social links, but actually never or rarely uses them, might be by far not the most influential node.

In a recent review [1], Holme and Saramäki present an overview about theoretical literature dealing with the treatment of temporal networks - but a precise framework is still missing. When researchers dealt with dynamic networks in the past, it was mostly from a *snapshot* perspective, i.e. considering the adjacency matrix of the underlying varying graph after certain time intervals and doing *static network analysis* on these snapshots. The problem with this approach is that it does not link the consecutive snapshots together to a time-global picture, but rather treats them as separate entities. In such a framework it is hard to study e.g. the question of possible *time-respecting paths* [1] in a precise analytical fashion. Also, it becomes hard to define centrality metrics that go beyond the obvious aggregation of the values obtained in the snapshots. What is temporal betweenness? How should temporal closeness be defined? Is it possible in the first place to give precise definitions of such, based on straight-forward analogies from static network theory, given the new degree of freedom? With this poster, we present a promising general framework to deal with such situations, called *time-explicit graphs (txg)*.

2. TIME-EXPLICIT GRAPHS

One of the main questions for the analysis of dynamic network topologies is the possible *information flow* (in a wide sense) from node to node. This consideration inspires the idea to represent the dynamic graph as a *static* network of flows, which we call time-explicit graphs[1]. This framework is best understood considering the example of figure 1. Start-

[1]We learned that a similar framework has been presented recently and independently by Kim and Anderson [2]

Summer School: Modeling and Analysis of Novel Mechanisms in Future Internet Applications 2011, Würzburg, Germany

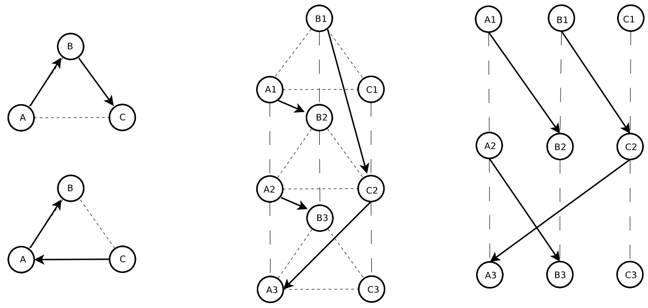

Figure 1: Transformation of a series of snapshots of dynamic graphs to the time-explicit graph.

ing with two consecutive snapshots of a network containing three nodes, it is possible to rewrite this as a network of flows as depicted in the middle figure. Here, the vertical dimension is the time dimension. As can be seen in the left figure, in the first time step it is possible to have information flow either from A to B or from B to C. It is this possible information flow within *one* time-step, which is encoded in the middle figure as the two connections from either node A in the first layer to node B in the second layer and from node B in the first layer to node C in the second layer. Since there is no structural information within the vertical layers, the middle figure can be safely transformed into the figure on the right. It is this kind of structure, which we call time-explicit graph. From the simple toy-model in figure 1 there are already a couple of interesting conclusions one can draw.

- Following the possible information flow, it is seen without any problem that in the dynamically changing graph there exists exactly one time-respecting path, which is B1-C2-A3. Assumaning that from one time-step to another information is either propagated or lost, this implies that only information starting at node B at time 1 will still be in the system at time 3, and will reside at node A.
- Node C might be considered an important ("central") node in the network, since it basically links the "past" to the "future". Node C is a "temporal short-cut". This is in stark contrast to the snapshot-picture of the left-figure, in which one would have argued that node B (in the upper) and node C (in the lower) are the most "central" nodes. Indeed, it is solely these two nodes which would have a betweenness-score $\mathcal{B} > 0$ in the snapshot-picture, hence neglecting the "temporal importance" of node C.

Since the time-explicit graph is a static network (to be precise it is a directed polytree), it is possible to use standard network-theoretic tools and measures to analyze it. In this sense, the first observation raised above can be addressed by calculating the number of connecting paths between the first time layer and the last time layer. If the timely distance is t, this can easily be done via checking for non-zero entries of the matrix $(A_{txg})^t$ with A_{txg} being the adjacency matrix of the time-explicit graph. For the purpose of general quantitative analysis, we provide a set of python scripts on the web page of the project, based on igraph for python. It is clear that with this framework we hence transform the additionally introduced time complexity into complexity of the size of the studied time-explicit graph. To be precise, given $T - 1$ consecutive snapshots of graph topologies, the time-explicit graph will have T layers of nodes. Since, for the case of a closed system, the number of nodes is the same in every layer, say n, the total number of nodes of the time-explicit graph is $N = n{\cdot}T$ and hence the size of the adjacency matrix A_{txg} is $N \times N$. Because of the polytree-like structure of the txg, it is a very sparse graph with at most $(T-1) \cdot n^2$ edges, as compared to $N^2 = T^2 n^2$. This sparsity is beneficial for storage and time-complexity of computations performed on the txg. Hence the txg-approach also provides a computationally promising approach to dynamic network studies.

3. POSSIBLE APPLICATIONS

Extending network analysis to a dynamic topology setting is not simply of theoretical interest, but has manifold practical applications. Especially in the area of socio-technical systems we expect such a framework to be useful. For example the question of ad-hoc networking, i.e. networking in highly volatile routing overlay networks, might benefit from such a modelling and analysis framework. Especially when the underlying dynamics (which is the sociological component) is known (or can be approximated), the framework presented here might help to design such systems in some optimal way. This also goes in-line with the recent effort to design massive distributed computing systems [5]. Also, the presented framework might be useful in improving our understanding of disease-spreading phenomena (e.g. malicious software) in topologically volatile environments.

4. FUTURE WORK

The work on this framework has just started and there is still a lot to explore about time-explicit graphs. First, we will continue studying the meaning of "classical" network metrics in the time-explicit graph picture, which is part of the bigger goal to establish useful definitions for "temporal centrality". Indeed, in a recent comment [4] and working paper [3] we propose exactly such a general definition. Second, there are many ways to extend and modify the time-explicit graphs, e.g. including "temporal self-loops" (which might account for memory on the node side), introducing periodic boundary conditions (i.e. repeating topology dynamics) or the study of open instead of closed systems (i.e. varying number of nodes per time-layer).

5. REFERENCES

[1] P. Holme and J. Saramaki. Temporal Networks. *arXiv:1108.1780v2*.

[2] H. Kim and R. Anderson. Temporal node centrality in complex networks. *Physical Review E*, 85(2):1–8, Feb. 2012.

[3] R. Pfitzner, I. Scholtes, C. J. Tessone, A. Garas, and F. Schweitzer. A general framework for temporal graph analysis. *Working-paper*, 2012.

[4] R. Pfitzner, I. Scholtes, C. J. Tessone, A. Garas, and F. Schweitzer. Comment to "Temporal node centrality in complex networks". *Working-paper*, 2012.

[5] I. Scholtes and C. J. Tessone. Organic Design of Massively Distributed Systems : A Complex Networks Perspective. *Informatik-Spektrum*, (Special Issue on Organic Computing):1–15, 2012.